

NoSQL Proof-of-Concept Proposal

Team name: PSG.LGD

Yichun Yan

Ziwei Jiang

Yifan Li

Weiqi Wang

September 15, 2019

Contents

1	Dota2 Game Replay Analysis	2
1.1	Data Preprocessing and Specification	2

1 Dota2 Game Replay Analysis

1.1 Data Preprocessing and Specification

As the raw data is completely unstructured `.dem.bz2` file, preliminary preprocessing must be done before we store the data in our database.

Firstly, we can decompress the `.bz2` file with `bzip2`, which will give us a `.dem` file:

```
$ bzip2 -d data.dem.bz2
```

Next, we can utilize [clarity](#), an open source Dota2 replay parser, to extract useful information from the `.dem` file. Based on our exploration and the [examples](#) provided by clarity, the following data will be available:

- Player name, id, team formation and hero choice
- Detailed log of the game, including a hero:
 - deals damage to another one
 - heals another one
 - receives/loses a buff/debuff
 - kills another one
 - uses his ability
 - uses an item
 - buys an item
 - receives/loses some gold
 - gains some XP
 - buys back (spending money in order to instantly re-spawn)
- Spawn/death of heros and NPCs
- Summary of each player's performance in the whole match, including:
 - Final level
 - Kills
 - Deaths
 - Assistance
 - Gold
 - Last hit
 - Deny

The following fields are critical to answering our Business Question, “???”. They will need to be cleansed and validated.